Response 1: Scientific Method in Social Policy Research is Not a Lost Cause

McKay, Stephen

DOI:
10.1017/S0047279410000589

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
McKay, S 2011, 'Response 1: Scientific Method in Social Policy Research is Not a Lost Cause', *Journal of Social Policy*, vol. 40, no. 1, pp. 21-29. https://doi.org/10.1017/S0047279410000589

Response 1: Scientific Method in Social Policy Research Is Not a Lost Cause

STEPHEN McKAY

Journal of Social Policy / Volume 40 / Issue 01 / January 2011, pp 21 - 29
DOI: 10.1017/S0047279410000589, Published online: 21 July 2010

Jnl Soc. Pol. (2011), 40, 1, 21–29 © Cambridge University Press 2010
doi:10.1017/S0047279410000589

Response 1: Scientific Method in Social Policy Research Is Not a Lost Cause

STEPHEN McKAY*

*School of Social Policy, University of Birmingham, Birmingham B15 2TT
email: s.d.mckay@bham.ac.uk

Introduction

In 'Generalisation and phronesis', Paul Spicker argues that the methodology of empirical research in social policy would be improved by drawing on Aristotle's concept of *phronesis* (φρόνησις) – which is generally translated as practical wisdom or prudence. He argues against versions of generalisation that rely on cause-and-effect, whether through deductive or inductive means. He maintains, however, that social policy research must still aim at making generalisations, for which he recommends the application of phronesis.

Social policy academics can be grateful that Spicker has chosen to initiate (or, might I say, to cause?) a debate on these important methodological issues. Our methodology is fundamental to what we do. However, in this response I argue that Spicker has failed to convince that social policy should abandon a cause-and-effect-led research strategy. If his views are taken in a weaker form – that it is best practice to reflect on experience and be aware of the key role of context in formulating and testing theories – then he is suggesting something that is already true in social policy research, and undeniably valuable.

Overview of the argument for using phronesis

There are a number of steps in Spicker's argument for phronetic social policy research. Some of the earlier sections of the article are not quite so relevant for the main argument, but provide supporting narrative. The key steps in the logic of the article seem to be as follows:

1. That social policy research, and particularly research for policy purposes, is based on the idea that research results may be generalised. This much is presumed true within the social sciences more generally.
2. That these generalisations are based on a 'cause-and-effect' view of the world. Such a view may be based within different research methodologies, whether they be deductive (theory-led), inductive (generalisation from observed data) or critical realist in approach.

3. That causal accounts may be of various kinds, ranging from a 'necessary connection' perspective, through to more probabilistic accounts, to narrative accounts.
4. That causal accounts of the world cannot be trusted, either on the grounds of philosophical inquiry or owing to the complexity of social phenomena:

   > And, in any case, causal accounts are not always helpful in formulating a policy response to given problems.

5. That social policy must still pursue generalisation as an objective, rather than fall into a relativist position on ontology.
6. That phronesis provides an *alternative* route to providing generalisations that may assist in understanding social policy problems, and one that is distinct to and different from relying on causal accounts of the world.
7. This then leads to a reflective section on how to distinguish good from bad generalisations. He advises that good versions must (a) begin from evidence not theory, (b) be cross-confirmed by examples in other fields, and (c) be 'made about processes in their social context'.

On the first point (1), it is certainly arguable whether research in social policy is always based on the idea of generalisability. Indeed, much research aims at quite localised understanding and interpretations of people's experiences. However, it is perfectly possible to consider, and to challenge, Spicker's arguments by focusing on the areas of research, especially policy-related research, where there is the more explicit aim of arriving at findings that may be generalised. Spicker's argument is actually directed at research for policy purposes.

The details of points (2) and (5) should also not detain us for long. The former is the target of Spicker's article, and the nub of the argument is really later in (6) when specific examples are shown as illustrative cases of phronesis rather than of standard cause-and-effect approaches. Point (5) is the argument that the social sciences must be able to generalise to make any progress. The alternative to seeking generalisable statements is to adopt a relativist view, where reality is constructed in people's minds such that generalisation is impossible. Some may wish to dispute this, but it is accepted as axiomatically true in this analysis that generalisation is the ultimate end of the research process, although that may not be true of every single research study conducted.

Overall, then, the key steps in the argument are stages (3), (4), (6), and (7), and we discuss each in turn in the next section.

The key steps in the argument

Parts (3) and (4): approaches to causation and their problems

Spicker identifies different approaches to causation in step (3). The standard required for causation to be a necessary connection (cf. Hume, 1739, 1748) fails

even in modern 'hard science' grappling with, *inter alia*, the uncertainty principle in quantum mechanics, and chaos theory. Complete determinism is simply no longer credible in any kind of science. If it is required that causes are both necessary and sufficient for their effects, it would be impossible to claim that 'smoking causes lung cancer', for instance – because not all smokers contract this disease, and not all those with this disease were smokers. Yet most will agree that it is a causal claim with a great deal of support, tantamount to proof. Hence the adoption of the so-called Bradford-Hill (1965) criteria for determining a causal association, at least in epidemiology, but also with resonance in other subjects based on observation rather than experiment.[1] These criteria include such notions as consistency and plausibility, which involve a degree of expert judgement and yet are clearly aiming at causal explanations.

Therefore perhaps the causal account that most will recognise in their empirical work is the probabilistic account of causation (Humphreys, 1989). We do not attempt to predict outcomes at the level of individuals, but to judge probabilities and thereby to consider outcomes at the level of larger groups.

One of the key claims made is in part (4) of the argument – that causal accounts may not be trusted. Spicker draws on philosophical claims that 'cause-and-effect' may not be trusted by quoting Hume's *Treatise of Human Nature* (1739). Hume is, of course, arguing against causation in general, rather than in the social sciences. Causation within natural science is also based on shaky philosophical foundations, according to this influential account. Yet, in this appeal to the authority of David Hume, no other philosophical critiques are offered despite the fact that philosophical work on causation clearly did not stop with Hume, but continues with contributions looking at counterfactuals (Lewis, 1973) and at probabilistic causation (Humphreys, 1989). The idea of entirely determined causation is no longer credible, but that does not mean that causation is itself undermined. Nor are the philosophical issues truly settled.

The main argument deployed to raise scepticism about causal accounts is that causal analysis is difficult because social phenomena are complex. Very little of the article is spent on this argument, but this is the pivot point of much of the whole logic of the piece. This 'argument from complexity' has been tackled in at least some of the literature on the philosophy of social science (McIntyre, 1996: 17, 56). Spicker does not really present much of an argument for this, and merely states the idea of complexity – overall, it would be helpful to consider the evidence for this, and its implications. Certainly, McIntyre's in-depth analysis of this argument reaches a different conclusion to that of Spicker.

Spicker goes on to claim that it is no defence of causal analysis that it provides useful prescriptions for policy. It is certainly true that knowing the nature of a phenomenon provides no direct guidance on what policy *should* do to tackle social problems. Policy-makers need to know, arguably, what will help jobseekers to find work, and what will help in assisting smokers to stop, rather than knowing

who is more likely to be unemployed, or what groups of people are the most likely to start to smoke. But that is to say that the wrong problem is often being studied by researchers, not that causal analysis itself is of no benefit. Spicker's person at the bottom of a well needs hypotheses on what approaches are feasible, and ideally evidence on whether trying to climb out has a greater chance of success than shouting for help.

Part (6): the alleged distinctiveness of selected phronetic generalisations in social policy

There have been a number of calls for a social science based more squarely on phronesis, and Spicker cites Flyvbjerg (2001). But in a similar vein we may also list the work of Stephen Toulmin (2001). He traced the separation in Western thought into two different strands of 'rationality' (scientific, certain) and 'reasonableness' (narrative, uncertain), and described how the latter only came to be viewed as superior in modern times. Spicker would find some common ground in this analysis, I think. MacIntyre (1985) also suggested a move towards Aristotelian approaches. These authors differ on key points, but all have made similar suggestions for tackling what they see as the difficulties and crises in social science research.

One of the key issues is how far Spicker's article is essentially applying the terms of the 'Flyvbjerg debate' within political science (Part 1 of Schram and Caterino, 2006) to social policy, or whether it is doing something different to, or beyond, that. The work of Flyvbjerg was influential in political science, and there are a number of responses from that discipline that may be used to critique the idea of a phronetic social science within social policy. For example, Laitin (2003) conducted a robust defence of scientific method in political science, contra Flyvbjerg. Furthermore, even if a phronetic approach has its merits, why must it be advanced as a candidate for a new dominant methodology which *replaces* existing approaches, rather than simply adding to the set of approaches at our disposal? (cf. Jackson, 2006).

Moreover, social policy research does not have the same kinds of problems that Flyvbjerg (2001) was addressing when he made a comparable argument about more quantitative social sciences (addressing disciplines apparently afflicted either by 'physics envy' or 'economics envy'). Most in social policy would agree that research should 'address problems that matter to groups in the local, national and global communities in which we live' and that 'we must effectively . . . communicate the results of our research to our fellow citizens and carefully listen to their feedback' (Flyvbjerg, 2005: 42). This may well contrast with some subjects where 'quantitative analysis, formal modelling and hard science dominate the leading journals and research institutions' (Kasza, 2006: 222), but is hardly the case for social policy journals in the UK or an imminent risk (see Smith, 2008: 327 for an analysis of Education, Sociology and Social Work). The

reverse, a lack of quantitative and statistical capacity, is more likely to be the case (Mills *et al.*, 2006).

As an example of phronesis, Spicker quotes the link between unemployment and sickness benefits. He states that this is different from other kinds of causal explanations, but does not provide an argument for this being so. The hypothesis that 'an increase in unemployment leads to an increase in claims for sickness (or disability) related benefits' seems as causal an hypothesis as any other, and subject to empirical testing (and may be made falsifiable) – Beatty *et al.* (2000) may be considered an example of such an approach. One might imagine there is then scope for academic judgement about how adequately to conceptualise measurements, time lags and the like when researching such a theory. Curiously, he states 'It is not an explanation of what happens; it is a description' – which begs the question of how a useful description cannot also invoke some kind of explanation.

Spicker, in fact, provides a set of examples, which are claimed to be non-causal, and all dependent on circumstances. In particular, a lot must hinge on how these studies differ, if at all, from the probabilistic model of causation outlined earlier. Two of the examples included are research into non-take-up of benefits (Craig, 1991) and the 'inverse care *law*' (Tudor Hart, 1971, my emphasis). Spicker argues that these are expressions of phronesis, and so cannot be said to be causal accounts or explanatory. In the next couple of paragraphs, I say a bit more about these studies, and critically evaluate whether or not they may be said to be causally based rather than examples of phronesis.

Research on take-up has been through a number of generations, with rigorous testing of different theoretical models and predictions. Kerr (1983) constructed an empirical test of a model based around a set of thresholds rather than trade-offs. More recent research has sought to model the take-up process as related to the 'costs of claiming' (Pudney *et al.*, 2004). In each case, there were well-formed hypotheses which were then subject to empirical testing. The analysis of incomplete take-up of social security benefits is now a mature sub-field of investigation. For instance, it is well-established that the higher the level of entitlement, the greater the rate of take-up Oorschot (1995).

The latter example of an inverse link between the need for health care and its provision is presented as a 'law' rather than as a finding based on particular circumstances (Tudor Hart, 1971). This law states that:

> The availability of good medical care tends to vary inversely with the need for it in the population served. This inverse care law operates more completely where medical care is most exposed to market forces, and less so where such exposure is reduced. (ibid: 405)

Tudor Hart marshalled a range of historical and statistical resources to substantiate his 'law'. He did note that definitive statistical proof was not possible,

because relevant data were not always available, but it seems part of his aim (and past studies provided considerable support).

It is argued that these kinds of statements are not universally applicable, but instead depend on the circumstances. But I would argue that no-one presenting a causal view of the world is unaware of the importance of context and circumstances. In arriving at conclusions, it is actually helpful (rather than undermining) to be able to look at different contexts, and incorporate the variation across different circumstances into any model. This may even be more formally described within the 'most different' nation approach to comparative study (Przeworski and Teune, 1970). Aristotle himself was interested in classifying political systems based on their differences.

Part (7): distinguishing good from bad generalisations

In some respects the last point is among the most interesting and revealing of the approach. Clearly, Spicker is somewhat troubled by the implications of where his reasonings have led him. Without a 'hard' datum of evidence within a framework of cause-and-effect against which to test hypotheses, how can we reject (as Spicker wishes to) or refute the view of the likes of Charles Murray (1984) concerning the 'underclass'? Once we move away from the standard deductive approach of formulating an hypothesis subject to empirical testing ('falsifiable'), who is to distinguish which kinds of experiential phronetic generalisations may be relied upon from those we may wish to reject?

At the close of the article, Spicker is clearly concerned about how to separate what he sees as good or valid generalisations (for example, selective policies have incomplete take-up) from bad generalisations (such as the cycle of deprivation). Here his language betrays a return to a reliance on universal principles. Having argued that generalisations cannot take the form of laws, but are inevitably limited to context and the particular, and to judgement, he declares that some generalisations are nevertheless *untrue*. This is a strong universal claim, rather than one designed to apply in particular circumstances or only in certain contexts. But if generalisation is always subject to the kind of context and particular circumstances argued for elsewhere in the article, how is it possible to claim a generalisation of this kind is untrue (assuming it is not internally contradictory)? Of course, the answer is that it isn't possible.

The advice of experienced practitioners is not always to be trusted. In *Baby and Child Care* (1946), the renowned child care expert Dr Benjamin Spock said 'it is preferable to accustom a baby to sleeping on his stomach'. Chalmers (2001: 322) described this as 'apparently rational and authoritative advice', but which 'led to thousands, if not tens of thousands, of avoidable cot deaths' (see Gilbert *et al.*, 2005) Changes in the advice given to parents roughly halved the rate of cot deaths. When there is a conflict between the advice of practical wisdom and reliable empirical evidence, as in this case, it is clear which must give way.

Conclusions

Let me end as I began. Paul Spicker has produced a stimulating analysis that I commend to those both within and outside social policy. This is an excellent opportunity to review some of our key methodological challenges. I happen to disagree with many of the arguments presented, for the reasons outlined above.

A focus on phronesis is presented as being about the particular rather than the universal – 'when variations in conditions lead to different outcomes, we qualify the generalisation'. But nothing in this would give concern for those pursuing a causally based methodology. Indeed, for constructing and testing theories there is nothing better than having different conditions in which to test. Comparative research is an exemplar of this, where different contexts are an advantage (indeed, a requirement) in testing for example theories about the development of the welfare state.

Nor must we assume that by following the 'standard' method of inquiry that we are somehow following a regimented rather than a creative path. Many insights into research are creative, not the result of following a uniform series of steps. In investigating the increase in the number of heart attacks after the Second World War, Jerry Morris (sometime co-author with Richard Titmuss) thought to compare bus drivers with bus conductors – similar class origins, but different levels of exercise and rates of heart disease, leading to theories about the connection between them (Morris *et al.*, 1953[2]). Advice from C. Wright Mills remains valid, 'Only by conversations in which experienced thinkers exchange information about their actual ways of working can a useful sense of method and theory be imparted to the beginning student' (1959: 1).

In a brief review, Shaw (2003: 244) argued that Flyvbjerg (2001) arrives at 'straightforward conclusions, but chooses to take an elaborate path to get there'. It is hard not to extend at least some of this sentiment to Spicker's own narrative. Social policy research is not dominated by a quantitative approach, nor are its scholars generally cut off in their ivory towers from the mainstream of society. Long may that remain so. The currents of the subject are probably more away from formal approaches and statistical methodologies and towards interpretative approaches. Having a plurality of methods is a strength of the discipline, and that includes a role for the scientific method and the testing of causal theories against evidence. We may construct a 'straw man' of causation, where effects must necessarily and always follow from their causes, but that approach to causation has long since been discredited. That leaves open a role for academic judgement, but still within a cause-and-effect framework.

My overall verdict is that Spicker is either wrong to recommend that we abandon a cause-led empirical strategy, or he is merely suggesting something that is already accepted practice, namely to reflect on experience and be aware of the key roles played by context. The latter is good advice, and it is helpful to be reminded of this from time to time.

Acknowledgement

I am grateful to Karen Rowlingson for cogent comments on an earlier draft. I am also very grateful to the editors of the *Journal of Social Policy* for asking me to write this response, and to two anonymous referees for helpful comments that improved the article.

Notes

1 There are some affinities between social policy and epidemiology, in that both must generally rely on observed patterns of behaviour and outcomes rather than controlled experiments. Some of the topics are also common, and Richard Titmuss' early work included the study of child mortality – *Birth, Poverty and Wealth: A Study of Infant Mortality* (1943). Of course, there is also the possibility of more negative linkages along a 'pathological' line.

2 For a recent interesting, popular account, see *Financial Times*, 11 September 2009, 'The man who invented exercise', by Simon Kuper.

References

Beatty, C., Fothergill, S. and MacMillan, R. (2000), 'A theory of employment, unemployment and sickness', *Regional Studies*, 34: 7, 617–30.

Bradford Hill, A. (1965), 'The environment and disease: association or causation?', *Proceedings of the Royal Society of Medicine*, 58: 295–300.

Chalmers, I. (2001), 'Invalid health information is potentially lethal', *British Medical Journal*, (21 April), 322: 7292, 998.

Craig, P. (1991), 'Costs and benefits', *Journal of Social Policy*, 20: 4, 537–65.

Flyvbjerg, B. (2001), *Making Social Science Matter*, Cambridge: Cambridge University Press.

Flyvbjerg, B. (2005), 'Making social science matter', *Foresight Europe*, October 2005 and March 2006, pp. 38–42, accessed 8 June 2010 at http://flyvbjerg.plan.aau.dk/Publications2006/ForesightNo2PRINT.pdf.

Gilbert, R., Salanti, G., Harden, M. and See, S. (2005), 'Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002', *International Journal of Epidemiology* [doi:10.1093/ije/dyi088]

Hume, D. (1739), *A Treatise of Human Nature*, Oxford: Oxford University Press.

Hume, D. (1748), *An Enquiry Concerning Human Understanding* (section VII), Oxford: Oxford University Press.

Humphreys, P. (1989), *The Chances of Explanation: Causal Explanations in the Social, Medical, and Physical Sciences*, Princeton, NJ: Princeton University Press.

Jackson, P. T. (2006), 'A statistician strikes out: in defense of genuine methodological diversity', in S. Schram and B. Caterino (eds.), *Making Political Science Matter: Debating Knowledge, Research, and Method*, New York: New York University Press, pp. 86–97.

Kasza, G. (2006), 'Unearthing the roots of hard science', in S. Schram and B. Caterino (eds.), *Making Political Science Matter: Debating Knowledge, Research, and Method*, New York: New York University Press, pp. 222–33.

Kerr, S. (1983), *Making Ends Meet: An Investigation into the Non-claiming of Supplementary Pensions*, London: Bedford Square Press.

Laitin, D. (2003), 'The perestroika challenge to social science', *Politics and Society*, 31: 1, 163–84 [reproduced in Schram and Caterino (2006)].

Lewis, D. (1973), 'Causation', *Journal of Philosophy*, 70: 556–67.

MacIntyre, A. (1985), *After Virtue*, London: Duckworth.

McIntyre, L. (1996), *Laws and Explanation in the Social Sciences: Defending a Science of Human Behaviour*, Oxford: Westview Press.

Mills, C. W. (1959), 'On intellectual craftsmanship', appendix to *The Sociological Imagination*, Oxford: Oxford University Press.

Mills, D., Jepson, A., Coxon, T., Easterby-Smith, M., Hawkins, P. and Spencer, J. (2006), *Demographic Review of the UK Social Sciences*, Swindon: ESRC.
Morris, J., Heady, J., Raffle, P., Roberts, C. and Parks, J (1953), 'Coronary heart disease and physical activity of work', *Lancet*, 2: 1053–7.
Murray, C. (1984), *Losing Ground*, New York: Basic Books.
Oorschot, W. van (1995), *Realizing Rights*, London: Avebury.
Przeworski, A. and Teune, H. (1970), *The Logic of Comparative Social Inquiry*, London: John Wiley & Sons.
Pudney, S., Hancock, R. and Sutherland, H. (2004), 'Simulating the reform of means-tested benefits with endogenous take-up and claim costs', ISER Working Papers No. 2004-04, ISER, University of Essex
Schram, S. and Caterino, B. (eds.) (2006), *Making Political Science Matter: Debating Knowledge, Research, and Method*, New York: New York University Press.
Shaw, M. (2003), 'Flyvbjerg, 2001, Making social science matter' [book review], *Progress in Human Geography*, 27: 2, 244.
Smith, E. (2008), 'Pitfalls and promises: using secondary data in educational research', *British Journal of Educational Studies*, 56: 3, 323–39.
Spock, B. (1946), *The Common Sense Book of Baby and Child Care*, New York: Duell, Sloan & Pearce.
Titmuss, R. M. (1943) *Birth, Poverty and Wealth: A Study of Infant Mortality*, London: Hamish Hamilton.
Toulmin, S. (2001), *Return to Reason*, London: Harvard University Press.
Tudor Hart, J. (1971), *The Lancet* 297: 7696 (27 February), 405–12; also reprinted in G. Smith, D. Dorling and M. Shaw (eds.) (2001), *Poverty, Inequality and Health in Britain* 1800–2000, Bristol: Policy Press.